Correlation Detection and an Operational Interpretation of the Rényi Mutual Information

Masahito Hayashi[1,2] and Marco Tomamichel[3,2]

[1] *Graduate School of Mathematics, Nagoya University, Furocho, Chikusaku, Nagoya, 464-860, Japan*

[2] *Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore*

[3] *School of Physics, The University of Sydney, Sydney 2006, Australia*

Recently, a variety of new measures of quantum Rényi mutual information and quantum Rényi conditional entropy have been proposed, and some of their mathematical properties explored. Here, we show that the Rényi mutual information attains operational meaning in the context of composite hypothesis testing, when the null hypothesis is a fixed bipartite state and the alternate hypothesis consists of all product states that share one marginal with the null hypothesis. This hypothesis testing problem occurs naturally in channel coding, where it corresponds to testing whether a state is the output of a given quantum channel or of a "useless" channel whose output is decoupled from the environment. Similarly, we establish an operational interpretation of Rényi conditional entropy by choosing an alternative hypothesis that consists of product states that are maximally mixed on one system. Specialized to classical probability distributions, our results establish an operational interpretation of Rényi mutual information and Rényi conditional entropy.

Full version available at `arxiv: 1408.6894`.

Introduction. In order to distill useful measures of Rényi mutual information and Rényi conditional entropy from a plethora of possible definitions, it is important to find out which definitions correspond to relevant operational quantities. For this purpose, let us consider how efficiently an arbitrary bipartite correlated state ρ_{AB} on systems A and B can be distinguished from product states when the marginal of ρ_{AB} on A is known to be ρ_A. This problem can be regarded as the problem of detecting correlations in the state ρ_{AB}. Formally, we consider the following binary *composite hypothesis testing* problem for n copies of such a state[1]:

Null Hypothesis: The state is $\rho_{AB}^{\otimes n}$.

Alternate Hypothesis: The state is of the form $\rho_A^{\otimes n} \otimes \sigma_{B^n}$ with σ_{B^n} any state on n copies of B.

This problem figures prominently when analyzing the converse to various channel coding questions in classical as well as quantum information processing.[2] There, the problem is specified by a description of a channel $\mathcal{E}_{A'\to B}$ and a bipartite state $\rho_{AA'}$ where the system A constitutes an environment of the channel, A' is the channel input, and B its output. We are given an unknown state on n copies of A and B and consider the following two hypotheses.

Null Hypothesis: The state is the output of n uses of the channel $\mathcal{E}_{A'\to B}$, namely the state is exactly $\rho_{AB}^{\otimes n}$ where $\rho_{AB} := \mathcal{E}_{A'\to B}[\rho_{AA'}]$.

Alternate Hypothesis: The state is the output of a "useless" channel and decoupled from the environment, namely it is of the form $\rho_A^{\otimes n} \otimes \sigma_{B^n}$ with σ_{B^n} any state on n copies of B.

---

[1] We want to consider the speed with which the probability that we erroneously support the state $\rho_{AB}^{\otimes n}$ when the actual state is a product state of the form $\rho_A^{\otimes n} \otimes \sigma_B^{\otimes n}$ under a constraint for the opposite error. As is explained later, this problem can be discussed as the Hoeffding bound and Stein's lemma under this formulation.

[2] There exists an intimate connection between quantum channel coding and binary hypothesis testing (see, e.g., [8]). This connection is particularly important when analyzing how much information can be transmitted with a single use of a quantum channel [11, 23] or when approximating how much information can be transmitted with finitely many uses of the channel [3, 22]. (See also [7, 19] for the classical case. In particular, Polyanskiy [18, Sec. II] discusses the classical special case of this hypothesis testing problem.)

A *hypothesis test* for this problem is a binary positive operator-valued measure $\{Q_{A^nB^n}, \mathbf{1}_{A^nB^n} - Q_{A^nB^n}\}$ on the n copies of the systems A and B, determined by an operator $0 \leq Q_{A^nB^n} \leq \mathbf{1}_{A^nB^n}$. If the operator $Q_{A^nB^n}$ "clicks" on our state, we conclude that the null hypothesis is correct, whereas otherwise we conclude that the alternate hypothesis is correct. The *error of the first kind*, $\alpha_n(Q_{A^nB^n})$, is defined as the probability with which we wrongly conclude that the alternate hypothesis is correct even if the state is $\rho_{AB}^{\otimes n}$, given by

$$\alpha_n(Q_{A^nB^n}) = \mathrm{tr}[\rho_{AB}^{\otimes n}(\mathbf{1}_{A^nB^n} - Q_{A^nB^n})]. \tag{1}$$

Conversely, *the error of the second kind*, $\beta_n(Q_{A^nB^n})$, is defined as the probability with which we wrongly conclude that the null hypothesis is correct even if the state is of the form $\rho_A^{\otimes n} \otimes \sigma_{B^n}$ for some σ_{B^n}, given by

$$\beta_n(Q_{A^nB^n}) = \max_{\sigma_{B^n}} \mathrm{tr}[\rho_A^{\otimes n} \otimes \sigma_{B^n}\, Q_{A^nB^n}], \tag{2}$$

where the maximum is taken over all states σ_{B^n} on n copies of B.

Main Results. The main contribution of this paper is an asymptotic analysis of the fundamental trade-off between these two errors as n goes to infinity. To investigate this trade-off, we ask the following questions: let us assume that our test is such that $\beta_n(Q_{A^nB^n}) \leq \exp(-nR)$, what is the minimum value of $\alpha_n(Q_{A^nB^n})$ we can achieve? The answer is different depending on whether R is smaller or larger than the *mutual information* between A and B, denoted $I(A\!:\!B)_\rho$. If $R < I(A\!:\!B)_\rho$, we show that the minimal error of the first kind vanishes exponentially fast in n. This implies a *quantum Stein's lemma* [9] for the above composite hypothesis testing problem.

More formally, we define

$$\hat{\alpha}_n(nR) = \min_{0 \leq Q_{A^nB^n} \leq \mathbf{1}} \left\{ \alpha_n(Q_{A^nB^n}; \rho_{AB}) \,\middle|\, \beta_n(Q_{A^nB^n}) \leq \exp(-nR) \right\} \tag{3}$$

and investigate the exact exponents with which this error vanishes as n goes to infinity, yielding a *quantum Hoeffding bound* [6, 15] for our composite hypothesis testing problem. We find that the exponents are determined by the *Rényi mutual information*, defined as

$$I_\alpha(A\!:\!B)_\rho = \min_{\sigma_B} D_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B), \qquad \text{for} \qquad \alpha \in (0,1), \tag{4}$$

where $D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1} \log \mathrm{tr}\left[\sigma^{\frac{1-\alpha}{2}} \rho^\alpha \sigma^{\frac{1-\alpha}{2}}\right]$ is the Rényi relative entropy first investigated by Petz (see, e.g. [17]) and the minimization is over all states σ_B on B. We obtain

$$\lim_{n\to\infty} \left\{ -\frac{1}{n} \log \hat{\alpha}_n(nR) \right\} = \sup_{s \in (0,1)} \left\{ \frac{1-s}{s} \big(I_\alpha(A\!:\!B)_\rho - R \big) \right\}. \tag{5}$$

On the other hand, if $R > I(A\!:\!B)_\rho$, we show that $\hat{\alpha}_n(R)$ must approach one exponentially fast in n. This implies the strong converse for quantum Stein's lemma [16] for our problem. We then find the exact exponents (also called *strong converse exponenents*, see [5, Ch. 3] and [12, 16]) with which the error of the first kind goes to one as n goes to infinity and we find that in our case the exponent is determined by the *sandwiched Rényi mutual information* [1, 4], given as

$$\widetilde{I}_\alpha(A\!:\!B)_\rho = \min_{\sigma_B} \widetilde{D}_\alpha(\rho_{AB} \| \rho_A \otimes \sigma_B), \qquad \text{for} \qquad \alpha > 1, \tag{6}$$

where $\widetilde{D}_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1} \log \mathrm{tr}\left[\left(\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha\right]$ is the (sandwiched) Rényi divergence [14, 24]. We obtain

$$\lim_{n\to\infty} \left\{ -\frac{1}{n} \log \left(1 - \hat{\alpha}_n(nR)\right) \right\} = \sup_{s>1} \left\{ \frac{s-1}{s} \left(R - \widetilde{I}_s(A:B)_\rho \right) \right\}. \tag{7}$$

Hence, we show that the above composite hypothesis testing problem yields an operational interpretation for different definitions of the Rényi mutual information for the two ranges of α, paralleling the observation in [12].

Finally, we also perform a second-order analysis for quantum Stein's lemma [10, 21] and show that the minimal error of the first kind converges to a constant if $\beta_n(Q_{A^nB^n}) \leq \exp(-nI(A:B)_\rho - \sqrt{n}r)$ for some $r \in \mathbb{R}$. Then, for any $r \in \mathbb{R}$, we have

$$\lim_{n\to\infty} \left\{\hat{\alpha}_n\big(nI(A:B)_\rho + \sqrt{n}\,r\big)\right\} = \Phi\left(\frac{r}{\sqrt{V(A:B)_\rho}}\right), \tag{8}$$

where Φ is the cumulative standard normal (Gaussian) distribution and

$$V(A:B)_\rho := \operatorname{tr}\left[\rho_{AB}\big(\log\rho_{AB} - \log\rho_A\otimes\rho_B - I(A:B)_\rho\big)^2\right]. \tag{9}$$

is the *mutual information variance*.

Conditional Entropy. Analogously, an operational interpretation for *conditional Rényi entropies* is established by considering the following binary hypotheses testing problem, which is motivated by the task of decoupling of quantum states. The problem is specified by a description of a state ρ_{AB}. Given an unknown state on A and B, consider the following two hypotheses:

Null Hypothesis: The state is the n-fold product of ρ_{AB}, namely $\rho_{AB}^{\otimes n}$.

Alternate Hypothesis: The state is uniform on A^n and decoupled form B^n, i.e. it is of the form $\pi_A^{\otimes n}\otimes\sigma_{B^n}$, where π_A is the fully mixed state on A.

The same analysis as above applied to this problem reveals that the exponents in the quantum Hoeffding bound are determined by the Rényi conditional entropies defined as [20]

$$H_\alpha^\uparrow(A|B)_\rho = -\min_{\sigma_B} D_\alpha(\rho_{AB}\|\mathbf{1}_A\otimes\sigma_B), \qquad \text{for} \qquad \alpha\in(0,1)\,, \tag{10}$$

and the strong converse exponents are determined by the *sandwiched* conditional Rényi entropies [14]

$$\widetilde{H}_\alpha^\uparrow(A|B)_\rho = -\min_{\sigma_B} \widetilde{D}_\alpha(\rho_{AB}\|\mathbf{1}_A\otimes\sigma_B), \qquad \text{for} \qquad \alpha>1\,. \tag{11}$$

Main Proof Ideas. The main ideas are quickly summarized as follows:

1. We show that there exists a test (a different one in the direct and converse regime, respectively) which works uniformly for all choices of σ_{B^n} by using some elementary tools from group representation theory. Various tools, including a new minimax theorem are derived in order to prove the quantum Hoeffding bound.
2. To investigate the strong converse exponents, we show that the sandwiched Rényi mutual information can be achieved by pinching and show various other properties, including that it is differentiable in α for all $\alpha \geq \frac{1}{2}$. Then, we use the Gärtner-Ellis theorem for the large deviation analysis of the correlated distributions that result from pinching.

Related Work. Complementary and concurrent to this work, Cooney *et al.* [2] investigated the strong converse exponents for a similar hypothesis testing problem when adaptive strategies are allowed — however, they did not treat the case of a composite alternate hypothesis and they also did not analyze the error exponents in the quantum Hoeffding bound.

Our proof of the strong converse exponents parallels the development in a very recent preprint by Mosonyi and Ogawa [13]. There, the authors consider correlated states and use the Gärtner-Ellis theorem of classical large deviation theory in order to investigate the asymptotic error exponents in the presence of correlations. Here, we are not interested in correlated states *per se*, but our proof technique based on pinching naturally leads us to a classical hypothesis testing problem with correlated distributions, for which the Gärtner-Ellis theorem again provides the right solution.

[1] S. Beigi. Sandwiched Rényi Divergence Satisfies Data Processing Inequality. *J. Math. Phys.*, 54(12):122202, June 2013. DOI: 10.1063/1.4838855.

[2] T. Cooney, M. Mosonyi, and M. M. Wilde. Strong Converse Exponents for a Quantum Channel Discrimination Problem and Quantum-Feedback-Assisted Communication. Aug. 2014. arXiv: 1408.3373.

[3] N. Datta, M. Tomamichel, and M. M. Wilde. Second-Order Coding Rates for Entanglement-Assisted Communication. May 2014. arXiv: 1405.1797.

[4] M. K. Gupta and M. M. Wilde. Multiplicativity of Completely Bounded p-Norms Implies a Strong Converse for Entanglement-Assisted Capacity. Oct. 2013. arXiv: 1310.7028.

[5] M. Hayashi. *Quantum Information — An Introduction*. Springer, 2006.

[6] M. Hayashi. Error Exponent in Asymmetric Quantum Hypothesis Testing and its Application to Classical-Quantum Channel Coding. *Phys. Rev. A*, 76(6):062301, Dec. 2007. DOI: 10.1103/PhysRevA.76.062301.

[7] M. Hayashi. Information Spectrum Approach to Second-Order Coding Rate in Channel Coding. *IEEE Trans. on Inf. Theory*, 55(11):4947–4966, Nov. 2009. DOI: 10.1109/TIT.2009.2030478.

[8] M. Hayashi and H. Nagaoka. General Formulas for Capacity of Classical-Quantum Channels. *IEEE Trans. on Inf. Theory*, 49(7):1753–1768, July 2003. DOI: 10.1109/TIT.2003.813556.

[9] F. Hiai and D. Petz. The Proper Formula for Relative Entropy and its Asymptotics in Quantum Probability. *Commun. Math. Phys.*, 143(1):99–114, Dec. 1991. DOI: 10.1007/BF02100287.

[10] K. Li. Second-Order Asymptotics for Quantum Hypothesis Testing. *Ann. Stat.*, 42(1):171–189, Feb. 2014. DOI: 10.1214/13-AOS1185.

[11] W. Matthews and S. Wehner. Finite blocklength converse bounds for quantum channels. Oct. 2012. arXiv: 1210.4722.

[12] M. Mosonyi and T. Ogawa. Quantum Hypothesis Testing and the Operational Interpretation of the Quantum Renyi Relative Entropies. Sept. 2013. arXiv: 1309.3228.

[13] M. Mosonyi and T. Ogawa. The strong converse rate of quantum hypothesis testing for correlated quantum states. July 2014. arXiv: 1407.3567.

[14] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel. On Quantum Rényi Entropies: A New Generalization and Some Properties. *J. Math. Phys.*, 54(12):122203, June 2013. DOI: 10.1063/1.4838856.

[15] H. Nagaoka. The Converse Part of The Theorem for Quantum Hoeffding Bound. Nov. 2006. arXiv: quant-ph/0611289.

[16] T. Ogawa and H. Nagaoka. Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Trans. on Inf. Theory*, 46(7):2428–2433, Nov. 2000. DOI: 10.1109/18.887855.

[17] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Springer, 1993.

[18] Y. Polyanskiy. Saddle Point in the Minimax Converse for Channel Coding. *IEEE Trans. on Inf. Theory*, 59(5):2576–2595, May 2013. DOI: 10.1109/TIT.2012.2236382.

[19] Y. Polyanskiy, H. V. Poor, and S. Verdú. Channel Coding Rate in the Finite Blocklength Regime. *IEEE Trans. on Inf. Theory*, 56(5):2307–2359, May 2010. DOI: 10.1109/TIT.2010.2043769.

[20] M. Tomamichel, M. Berta, and M. Hayashi. Relating different quantum generalizations of the conditional Rényi entropy. *J. Math. Phys.*, 55(8):082206, Aug. 2014. DOI: 10.1063/1.4892761.

[21] M. Tomamichel and M. Hayashi. A Hierarchy of Information Quantities for Finite Block Length Analysis of Quantum Tasks. *IEEE Trans. on Inf. Theory*, 59(11):7693–7710, Nov. 2013. DOI: 10.1109/TIT.2013.2276628.

[22] M. Tomamichel and V. Y. F. Tan. On the Gaussian Approximation for the Classical Capacity of Quantum Channels. Aug. 2014. arXiv: 1308.6503.

[23] L. Wang and R. Renner. One-Shot Classical-Quantum Capacity and Hypothesis Testing. *Phys. Rev. Lett.*, 108(20), May 2012. DOI: 10.1103/PhysRevLett.108.200501.

[24] M. M. Wilde, A. Winter, and D. Yang. Strong Converse for the Classical Capacity of Entanglement-Breaking and Hadamard Channels via a Sandwiched Rényi Relative Entropy. *Comm. Math. Phys.*, 331(2):593–622, July 2014. DOI: 10.1007/s00220-014-2122-x.